UNIVERSITÉ DE MONTRÉAL

CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

A COMPARISON OF TWO METHODS FOR TACTICAL PLANNING IN RAIL FREIGHT TRANSPORTATION

Teodor Gabriel Crainic

Paper presented at the "10th Triennial Conference on Operations Research, IFORS'85, Washington, D.C., August 6-10 1984.

Ce rapport a été publié grâce à une subvention du fonds F.C.A.C. pour l'aide et le soutien à la recherche.

Université de Montréal
Centre de recherche sur les transports - Publication #335-A
Février 1984

Dépôt légal - 4e trimestre 1984

RÉSUMÉ

Dans cet article, nous comparons les résultats et performances d'un outil basé sur des modèles et techniques de recherche opérationnelle à ceux obtenus à l'aide des outils (système de simulation déterministe) présentement utilisés au Canadien National pour la planification tactique du transport des marchandises. Nous examinons les différences observées et les possibilités offertes par un outil de planification intégrant les deux approches.

ABSTRACT

We examine two different methods which may be used to assist the tactical planning process in rail freight transportation : a deterministic simulation method based on decision-making input from the planner and a network optimization model. The first has been designed at a large Canadian company, while the second has universitary origins. We compare the two methods and the characteristics of the solutions obtained and discuss the possibilities offered by a method combining the two approaches.

1. INTRODUCTION

A freight rail transportation system is a complex entity, larger than the sum of its parts and displaying intricate relationships and trade-offs between the policies affecting its different components. It is also a system where changes to the physical structure may take years to perform. It is therefore imperative to be able to forecast such changes and analyze their effects on the operations and performance of the system. Scenario evaluation is thus an essential part of the planning process of a railway company.

These characteristics are reflected in the complexity of the planning process, especially at the level of the medium (tactic) to long term (strategic) planning and good, fast and efficient tools are needed to assist the planner in his task.

In this paper we present two different approaches at building such tools. The first has been designed and implemented at the Canadian National Railways (C.N. Rail), a large Canadian company and it is based on a deterministic, interactive simulation model. The second is based on a network optimization model and has been developed at the Centre de recherche sur les transports of the Université de Montréal, with the support of C.N. Rail.

We begin by rapidly reviewing the problems associated with tactical planning in rail freight transportation. The two methods are then described and their main characteristics are emphasized. The following section contains results and comparisons between the two approaches and we conclude with a discussion of the perspectives opened by a possible combination of the two methods.

2. THE PROBLEM

The physical framework of any large railroad company consists of a complex network of small stations and large terminals interconnected by main and secondary lines. At any given time, a large number of cars (more than 100,000 C.N. and foreign cars, in the case of the C.N. Rail System) are in the network, and each may have an independant objective.

The cars, however, move in convoys called trains. In addition, cars are generally grouped into blocks, a block being a group of cars, with possibly different final destinations, arbitrarily considered as a single unit (for handling purposes) from the yard where it is formed till its destination, where its component cars are to be resorted. Rail companies use blocks as a mean to take advantage of some economies of scale and thus decrease the number of car manipulations in yards.

To monitor the individual car and train movements, rail companies use a number of detection and identification devices in connection with computer based systems, also used to collect data on those movements, as well as on train and block composition, yard workloads, etc.

To control the activities related to these individual movements, on the lines and through the yards, rail companies establish a series of operational policies, dynamic in nature, which deal, in great detail, with the local operations : train time-tables, track priority rules, dispatching policies in yards, engine and empty car scheduling and so on.

All these policies follow, however, the guidelines established at a higher level of decision-making, where general, network-wide operating strategies are generated : the tactical (medium-term) planning level. At this level of decision making, planners are mainly concerned with :

. classification policies : at each yard, what blocks to build (what are the yards for which blocks may be formed) and what cars to put into each block? How to distribute the workload among the yards of the system to fully use their capacities and avoid congestion?
. make-up policies : what blocks may be carried by each train?
. train routing policy, or how to establish the design of the service network and the level of service : what train services to run, on what routes and with what frequencies?
. traffic routing policy : determine for each origine-destination pair (possibly by commodity also) the routing pattern for its traffic : the train sequence used to travel and the yards where the cars are to be classified.

These policies have network-wide impacts and are strongly interconnected in both their economic aspects and their space-time dimensions. Trade-offs between the operating costs implied by a certain policy and its impact on the quality of the service (as reflected, in most cases, in the delays imposed on both freight and trains) have to be considered. The same trade-offs have also to be considered when the relationships between the four mentioned policies are examined.

The purpose of tactical planning is to specify an efficient and rational allocation of existing resources to improve the performance of the whole system. Day to day operations are not analyzed at this level of planning and decisions are sensitive only to broad variations in system parameters and data such as seasonal changes in traffic demand or modifications to the physical lay-out of the network.

The planner must then be provided with a tool which allows him to analyze, from a network point of view, the effects on the various elements of the rail system of changes in the configuration and characteristics of

the physical network (such as new or improved installations or the abandonment of lines), in external factors (such as energy prices or safety regulations), or in the global operating philosophy of the company (priority to intermodal COFC/TOFC transportation, for example). Thus this planning tool may be used for the tactical planning of the next season or to answer "what if" questions raised during strategic planning.

The tool must be flexible enough to facilitate the analysis of different scenarios and manageable enough to allow to experiment many strategies in a reasonable period of time. It may contain evaluation modules to analyze various attributes (capacity, congestion, costs) of individual units (yards and lines) but it must treat the system as a whole, at the network level. This also implies the capability to aggregate the network and the data according to the level of detail of the analysis.

We present, in the next two sections, two different approaches toward such a planning tool.

3. THE "CANAT" SYSTEM

We first consider the CANAT system, designed, developed and implemented at C.N. Rail with the goal of assisting transportation planners to measure the effects of changes in freight traffic on yards and lines from the perspective of the entire network [1]. This system, called CANAT for Computer-Assisted Network Analysis Tool, was designed as a macro tool for medium-to-long term planning, with an emphasis on the study of yard workloads.

Schematically, CANAT is a modular, interactive, deterministic simulation system and it requires the decision-making input of the planner about the rail system operations, capacities and structure.

CANAT is composed of a number of discrete steps, grouped into, currently, three modules (additional modules are in development).

The first module, Module 0, is used to generate the traffic forecast data, from historical car-trip data and from the long term forecasts established by the Marketing Department. These traffic forecasts are then reduced and organized on a quarterly basis by traffic type and origin-destination zones.

Module 1 is used to select the traffic data and to define the network for the study. For CANAT use, the more than 5000 original C.N. Rail stations are reduced to approximately 650 zones. Each zone represents a large terminal or an aggregation of a group of small yards, a particular line (with, possibly, a range of small stations) or a junction point. Then, for a particular study, a network of about 100 yards and 150 lines (the traffic demand originating from or with destination in a line zone is either eliminated or aggregated into a single pick- up/delivery point) is selected, boundary nodes are defined and the traffic data is reduced accordingly.

The third module is used to define the traffic movements to satisfy the demand and is made up of two parts. Each of the elements of the network, yards and lines, may have a number (from 5 to 30) of capacity-related parameters, such as yard capacity in number of trains originating/ending/working at the yard, maximum train length, line length, tonnage limits and so on. The first step of Module 2 is then used to fix these parameters and some other system elements such as standard routes for trains.

The second stage is used to interactively create trains to satisfy the demands for traffic movements and to check yard workloads. The traffic must be moved such as not to dramatically exceed the capacities defined in the previous stage. Trains are defined in terms of train class, speed

class, pick-up/set-off stations, traffic blocks carried and routes. The system may display informations about the status of the traffic demand, yard and line workloads, train blocks and so on. The user may create, modify or delete trains until satisfactory results are achieved and all traffic is successfully transported. A wide range of reports complete the system.

4. THE OPTIMIZATION METHOD

Our approach is to regard the tactical planning process for rail freight transportation as an optimization problem that we can model and solve by mathematical programming techniques.

The formulation is that of a network optimization model which aims to integrate the relationships and trade-offs between operating policies in train and traffic routing and classification and make-up policies in yards, at a tactical planning level. The purpose of the model is not to specify a detailed schedule but rather to generate economically sound global operating strategies ensuring a good level of service, in terms of transport delays and reliability.

A complete description of the formulation may be found in [3]. Here we only present the general modelling framework.

The physical network is aggregated following the same general rules as in CANAT. The line capacity-related parameters that we have retained include the line type (single, double), length, number of sidings, plus some train operational parameters such as a minimum headway for each train type. For each yard, one has to define an <u>average service time</u> for the classification operation. Presently, it is estimated from the capacity (maximum number of cars that may be classified during the planning period,

usually a week) of the yard as the mean through-time of a car being classified at that yard.

The traffic demand is specified for each traffic-class, defined as an origin-destination-type of commodity combination for which a positive number of cars (or any equivalent measure of tonnage) has to be moved.

Based on the physical network, the service network specifies the set of feasible routes on which train services may be operated. A train service is characterized by its type (defined by a speed-priority combination) and route (origin, destination, intermediate stops) in the physical network. The frequency of the service is defined as the number of trains operated on its route during the planning period and will be determined by the model.

An itinerary completely specifies a feasible journey for the traffic of a given traffic-class : the service path followed in the service network and the sequence of (re)classification yards along this path. The distribution of the traffic of each traffic-class among its itineraries will also be determined by the model.

The frequencies and itineraries are the central elements of the formulation, because fixing the values of the frequencies determine the services to be offered and the level of service, while by selecting for each traffic-class its best traffic distribution (for that level of service) between its itineraries, one solves the traffic routing problem and also determines network-wide classification and make-up strategies.

Schematically the model may be written as follows :

$$\text{Minimize} \quad \sum_m \sum_k Z_k^m (X,T) + \sum_h [Y_h(T) + P_h(X,T)]$$

$$\text{subject to } \sum_k x_k^m = d^m \qquad \text{all } m$$

$$x_k^m > 0 \qquad \text{all } m, \text{ all } k$$

$$t_h > 0 \text{ and integer} \qquad \text{all } h$$

where h, m, k index respectively services, traffic-classes, itineraries for each traffic-class and

x_k^m : quantity of the traffic of the traffic-class m using the itinerary k

X : vector of x_k^m for all traffic-classes, all itineraries

t_h: frequency of the service h

T : vector of frequencies

Z_k^m : total cost, operations and delays, for the itinerary k of the traffic-class m

Y_h : total cost, operations and delays, for the train service h

P_h : penalty cost, for the service h if the number of cars assigned to it exceeds the capacity of the service.

It is in the objective function that the relationships and trade-offs are considered. One important feature is the consideration of the delays incurred by trains and traffic due to congestion and operational policies in yards and on the lines. The delays are converted into costs via unitary delay costs defined for each type of service and each traffic-class. The major elements of the objective function are :

. Classification delays (for cars) : by means of queueing models (presently a M/M/1 model : Poisson arrival, exponential service, one server) expressing the mean classification delay as a function of the yard characteristics (the yard mean service time) and the total yard workload. For each yard, the number of admissible classification-destinations (and thus of blocks) is controlled via a penalty on small blocks.

. Handling costs for cars and trains on the road and in the yards.
. Over-the-road delays to train and cars. The mean delay is computed as a function of the characteristics of the link and the train congestion (by service type) on the link. We use formulae calibrated at C.N. Rail [5].
. Frequency delays, counting the time cars have to wait at yards until the designated service is available.

The resulting mixed integer, nonlinear, multiflow, multimode problem is solved by a decomposition algorithm working alternatively on the following subproblems :

. For a fixed traffic distribution, modify (if possible, stop otherwise) the frequencies to improve the global objective. A method of finite differences is used to determine the frequencies to modify.
. For a fixed level of service, determine an optimal traffic distribution. The problem is solved by a decomposition procedure which cyclically considers each traffic-class, all others variables being kept at their present values. For each subproblem, a new itinerary is first generated by shortest-path-like techniques, followed by the optimization of the traffic assignment done by descent methods.

The algorithm and its performance are analyzed in more details in [4]. Considering the size and complexity of the problems it proved to be a very effective algorithm. Some figures supporting this claim are also shown in Table 1 in the following section.

5. HOW THE METHODS COMPARE

Let's begin by summarizing the main characteristics of the two methods we just described. Il will help explain the observed differences in the behaviour of the methods and in the results obtained.

CANAT uses no explicit objective whatsoever, neither does it consider congestion or other time related functions. Neither operating costs, nor delay costs are considered and no automatic traffic assignment is performed. The planner has to specify each train and its workload, in this way effectively performing the traffic assignment and establishing the yard classification and make-up policies. He uses his judgment and experience and help from the status reports generated by the system. In this way, a detailed simulation of operations may be included into the system and the control of the planner over the form of the final results is total and continuous. The system however is unable to suggest new alternatives and the time necessary to build a solution is quite long (of the order of several days) which may limit the number of different strategies that may be examined.

The optimization approach do explicitely consider the delays in yards and over the roads and the costs attached to them and combines them with the operating costs (and some capacity-related penalties) in a global objective. Note that the formulation is flexible enough to allow the replacement of some or all cost and delay models presently included with other, more pertinent formulae. The model generates and evaluates a high number of operating strategies until it finds "the" best over-all strategy with respect to the general objective. At that moment, one obtains the services to be run, their frequencies and make-up, the traffic routing for each traffic class and the workload for each yard together with its classification policy. The proposed algorithm is quite efficient and may be decomposed into several procedures that may be used individually to do a traffic assignment for example, or to analyze the length of some train or the frequency of some service and so on. Of course, this method requires a large amount of data, some of which may not be readily available, such as

yard mean service times or, especially, delay costs by commodity. Also, being an optimization procedure, an important amount of detail and realism may be lost. Of course, details might be included but then the flexibility and rapidity will be lost. Therefore, a compromise has to be found.

When comparing the two methods, it rapidly becomes clear that the optimization approach is much faster but probably less precise. But, the comparison, in more detail, of the results generated by the two methods proved to be a more difficult task.

First, as emphasized by the characteristics we just mentioned, it is not possible to directly compare the two solutions on the basis of some objective function/economic criteria.

Second, due to technical differences (derived from the differences in the scope and methods of the two systems) in the outputs and the input requirements of the two computer programmes, we experienced significant difficulties in transfering the results of one system to the other. In fact, the only transfer that proved practically possible, was that of the service network generated through the use of CANAT.

We then conducted three types of experiments :

i) Generate the optimal traffic assignment using the CANAT service network and frequencies. We call this the "CANAT solution".

ii) Optimize using the CANAT service network.

iii) Optimize using a service network that we have generated.

We used two sets of data supplied by C.N. Rail, together with the service networks obtained by CANAT on them, as well as the passenger trains in operation on the network and estimations for the other parameters.

The first data set contained traffic demand forecasts for the year 1990 for the western part of the C.N. Rail's network. We had 735 traffic-

classes (after aggregations and counting only the traffic-classes with at least 10 cars/week ; in this way, we had over 95% of the total demand) and a network of 71 yards, 36 junctions and 95 links. The CANAT service network had 191 services and we generated a service network of 399 services. The three experiments on this data set will be refered to as Experiment 1, 2 and 3 respectively.

The second data set represented traffic demand forecasts for the year 1991 for the B.C. North Line, a subregion of the previous network on which an extensive study has been conducted at C.N. Rail. We had a network of 18 yards, 6 delivery/pick-up stations, 53 junctions and 41 links and 281 traffic-classes (capturing 99.32% of the total traffic demand). The CANAT service network had 47 services, while ours had 44 services. On this data set were conducted the Experiments 4, 5 and 6 respectively.

Some typical time-related figures for the resolution of the six problems are shown in the first table. They are in Cyber 173-equivalent C.P.U. seconds and the variations are explained by, among others, the size of the networks and of the service networks and the number of traffic-classes.

	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6
frequency modification procedure	---	40	133	---	1.1	0.9
traffic distribution	392	596	2277	76	105	87
itinerary generation	245	266	519	20	22	25
optimization	129	277	1443	37	55	34
total time	480	774	2518	91	126	108

Table 1 : Typical execution times for the various experiments

Table 2 shows the improvements obtained in the values of some performance parameters by the optimization procedure with respect to the respective CANAT solutions. Tables 3 and 4 display the level of service and the composition of the service networks obtained on the two data sets by the various experiments. Table 5 shows, for the three experiments conducted on the second data set, the distribution of the number of trains with respect to the length of the routes, while Table 6 contains, for the same problems, the distribution of the number of traffic classes according to the number of itineraries used.

system performance measures	Exp 2	Exp 3	Exp 5	Exp 6
total cost	5.48	1.01	6.90	10.43
ton-miles	0.37	- 0.39	2.65	2.06
car-miles	0.33	- 0.38	1.75	2.29
car costs				
delay	0.48	- 1.90	- 1.73	2.53
operation	0.28	- 1.20	2.02	1.90
total	0.31	- 2.92	1.35	2.00
train-miles	12.82	15.58	11.89	9.41
train costs				
delay	11.93	5.44	13.33	19.90
operation	11.55	5.76	12.48	18.97
total	11.63	5.69	12.67	19.17

Table 2 - Improvements (in %) in some system performance measures with respect to the CANAT solutions

type of service	Exp 1		Exp 2		Exp 3	
	#services	#trains	#services	#trains	#services	#trains
normal	177	1178	121	807	110	497
rapid	14	105	11	122	106	667
total	191	1285	132	929	216	1164

Table 3 - Service levels determined for the first set of data

type of service	Exp 4		Exp 5		Exp 6	
	#services	#trains	#services	#trains	#services	#trains
unit	4	68	4	68	4	68
normal	40	863	38	859	32	1080
rapid	3	49	2	16	0	0
total	47	980	44	943	36	1148

Table 4 - Service levels determined for the second set of data

length of the route in miles	Number of trains Exp 4	Exp 5	Exp 6
less than 200	79	52	36
less than 300	22	29	28
less than 400	170	150	278
less than 500	355	350	311
less than 1000	131	116	314
less than 2000	93	102	181
over 2000	157	144	0

Table 5 - Train distribution according with the route length

number of used itineraries per traffic-class	Number of traffic-classes Exp 4	Exp 5	Exp 6
1	255	236	270
2	28	15	13
3	11	7	2
4	12	7	0
5	9	12	0
6	0	8	0
total number of itineraries	407	423	304

Table 6 - Distribution of the number of traffic-classes according to the number of itineraries used

As these figures amply demonstrate, our method compares very advantageously with the CANAT system both in the speed of resolution and the quality of the solutions obtained.

The sixth experiment examplifies extremely well the differences between the two methods. In the CANAT approach, the main study was concentrated on the B.C. North Line, and thus all other traffic was moved by trains generated without much care. Our method however, had to optimize the transportation of all the traffic and thus achieved some very impressive improvements.

But, even for the other experiments, we obtained significant improvements. This is due to the characteristics of the optimization procedure which took full advantage of the economies of scale implied by larger blocks and full trains. The system also explored the possibility of using alternate routes for certain services and choose to operate, on certain services, shorter, more frequent trains to shorten the frequency delays and so improve the quality of service (and the overall objective).

One also notes that due to the cost structure C.N. Rail supplied us with, train costs (operation and delay) were much more significant than car costs. Consequently, the model optimized the rail system trying to save as much as posible on train service related costs, even if sometimes that meant increasing the car-costs.

There was a price to be paid for these performances, however, and it was the loss of some operational details. Thus some traffic-classes were put together in a block or a train in contradiction with some policy or operational rules.

Of course, these problems may be corrected, either by imposing more explicit or implicit constraints on the model and accepting the consequent loss of flexibility and efficiency, or by using a system in which the planner takes the solution as input and interactively modifies it to arrive at a valid policy. It is this second option that, in our opinion, has the better chances of success.

We see several possible utilization modes for our method when integrated in a comprehensive planning tool. In an exploratory mode, the system could be used to seek out possible operating strategies to respond to various scenarios developed during the strategic planning process. Costs and strategies could be produced and analyzed very rapidly with much less effort than currently. The system could also be used in an operating mode to help produce the set of policies to be implemented and used in the next planning period. The manager could interact with the system to modify a solution or to choose from several different solutions with basically the same objective function value. He could also modularly use the various procedures of the algorithm to analyze different parts of the system, such as the optimal traffic distribution for a given level of service or the train length on some of the routes offered.

Thus a system built upon our method may prove to be a valuable tool for the medium to long term management of a rail freight transportation system.

CONCLUSION

We have presented two different systems aimed at assisting the medium to long term planning process in rail freight transportation : the deterministic simulation approach presently in use at C.N. Rail and an optimization method presenting a comprehensive and flexible model together with an efficient algorithm. The optimization approach was much faster and generated very interesting results, suggesting that it might advantageously complement the methods presently in use.

ACKNOWLEDGEMENTS

The author wishes to take this opportunity to thank all those at C. N. Rail who have, at different degrees, contributed to this project : Paul N. Belshaw, W.J. Moore-Ede, Robert Feaney, James F.R. Gussow, P. Dirienzo, A. Stachursky and France Bourdeau.

REFERENCES

[1] Guide to Using CANAT, Exploitation Department, Canadian Railways (1982)

Crainic, T.G., Un modèle de planification tactique pour le transport ferroviaire des marchandises, PH. D. Thesis, publication #247, Centre de recherche sur les transports, Université de Montréal. (1982).

[3] Crainic, T.G., Ferland, J.A., Rousseau, J.-M., A Tactical Planning Model for Rail Freight Transportation, Transportation Science 18 (1984)

[4] Crainic, T.G., Rousseau, J.-M., Multimode, Multicommodity Freight Transportation : A General Modelling and Algorithmic Framework for the Service Network Design, publication #332, Centre de recherche sur les transports, Université de Montréal (1984).

[5] Hanks, N.A.R., Klein, R., Kuran, R.B., Moore-Ede, W.J., Elements of Railway Route Capacity Analysis, report Exploitation Department, Canadian National Railways (1974).